

Online Research @ Cardiff

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository: <https://orca.cardiff.ac.uk/id/eprint/104737/>

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

Pass, Daniel Antony ORCID: <https://orcid.org/0000-0003-2799-8432>, Sornay, Emily ORCID: <https://orcid.org/0000-0003-0720-2660>, Marchbank, Angela, Crawford, Margaret R., Paszkiewicz, Konrad, Kent, Nicholas A. ORCID: <https://orcid.org/0000-0002-4114-1307> and Murray, James A. H. ORCID: <https://orcid.org/0000-0002-2282-3839> 2017. Genome-wide chromatin mapping with size resolution reveals a dynamic sub-nucleosomal landscape in Arabidopsis. PLoS Genetics 13 (9) , e1006988. 10.1371/journal.pgen.1006988
file

Publishers page: <http://dx.doi.org/10.1371/journal.pgen.1006988>
<<http://dx.doi.org/10.1371/journal.pgen.1006988>>

Please note:

Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies.

See

<http://orca.cf.ac.uk/policies.html> for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.



RESEARCH ARTICLE

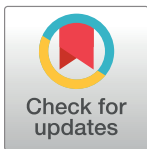
Genome-wide chromatin mapping with size resolution reveals a dynamic sub-nucleosomal landscape in *Arabidopsis*

Daniel Antony Pass^{1☯*}, Emily Sornay^{1☯}, Angela Marchbank¹, Margaret R. Crawford², Konrad Paszkiewicz³, Nicholas A. Kent¹, James A. H. Murray¹

1 Cardiff School of Biosciences, Cardiff University, Cardiff, Wales, United Kingdom, **2** Genome Centre, University of Sussex, Sussex House, Falmer, Brighton, United Kingdom, **3** Geoffrey Pope Building, University of Exeter, Stocker Road, Exeter, United Kingdom

☯ These authors contributed equally to this work.

* passda@cardiff.ac.uk



OPEN ACCESS

Citation: Pass DA, Sornay E, Marchbank A, Crawford MR, Paszkiewicz K, Kent NA, et al. (2017) Genome-wide chromatin mapping with size resolution reveals a dynamic sub-nucleosomal landscape in *Arabidopsis*. PLoS Genet 13(9): e1006988. <https://doi.org/10.1371/journal.pgen.1006988>

Editor: Ortrun Mittelsten Scheid, Gregor Mendel Institute of Molecular Plant Biology, AUSTRIA

Received: March 14, 2017

Accepted: August 21, 2017

Published: September 13, 2017

Copyright: © 2017 Pass et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All data are available from the ncbi geo database (accession number: GSE94377).

Funding: This work was supported by a Biotechnology and Biological Sciences Research Council grant BB/L009358/1 to JAHM and NAK. Sequence data was produced at the Exeter Sequencing Service and Computational core facilities at the University of Exeter (Medical Research Council Clinical Infrastructure award

Abstract

All eukaryotic genomes are packaged as chromatin, with DNA interlaced with both regularly patterned nucleosomes and sub-nucleosomal-sized protein structures such as mobile and labile transcription factors (TF) and initiation complexes, together forming a dynamic chromatin landscape. Whilst details of nucleosome position in *Arabidopsis* have been previously analysed, there is less understanding of their relationship to more dynamic sub-nucleosomal particles (subNSPs) defined as protected regions shorter than the ~150bp typical of nucleosomes. The genome-wide profile of these subNSPs has not been previously analysed in plants and this study investigates the relationship of dynamic bound particles with transcriptional control. Here we combine differential micrococcal nuclease (MNase) digestion and a modified paired-end sequencing protocol to reveal the chromatin structure landscape of *Arabidopsis* cells across a wide particle size range. Linking this data to RNAseq expression analysis provides detailed insight into the relationship of identified DNA-bound particles with transcriptional activity. The use of differential digestion reveals sensitive positions, including a labile -1 nucleosome positioned upstream of the transcription start site (TSS) of active genes. We investigated the response of the chromatin landscape to changes in environmental conditions using light and dark growth, given the large transcriptional changes resulting from this simple alteration. The resulting shifts in the suites of expressed and repressed genes show little correspondence to changes in nucleosome positioning, but led to significant alterations in the profile of subNSPs upstream of TSS both globally and locally. We examined previously mapped positions for the TFs PIF3, PIF4 and CCA1, which regulate light responses, and found that changes in subNSPs co-localized with these binding sites. This small particle structure is detected only under low levels of MNase digestion and is lost on more complete digestion of chromatin to nucleosomes. We conclude that wide-spectrum analysis of the *Arabidopsis* genome by differential MNase digestion allows detection of sensitive features hereto obscured, and the comparisons between genome-wide subNSP profiles reveals dynamic changes in their distribution, particularly at distinct genomic locations (i.e. 5'UTRs). The method here employed allows insight into the complex influence of

(MR/M008924/1). The Wellcome Trust Institutional Strategic Support Fund (WT097835MF), Wellcome Trust Multi User Equipment Award (WT101650MA) and BBSRC LOLA award (BB/K003240/1) contributed to library preparations and sequencing. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

genetic and extrinsic factors in modifying the sub-nucleosomal landscape in association with transcriptional changes.

Author summary

DNA is packaged by proteins into chromatin, of which the fundamental unit is a complex of histone proteins that wraps ~150bp of DNA into a nucleosome. Digestion of chromatin with enzymes such as micrococcal nuclease cuts the DNA between the protein particles, and by sequencing the cut sites, they can be mapped across the entire genome. Whilst nucleosomes are the most stable feature, less intensive digestion reveals a wider range of protein particles bound to DNA, in particular small sub-nucleosomal particles located most frequently upstream of genes. Here we show that these techniques can be used in the plant *Arabidopsis* to map the chromatin landscape of the range of particles sizes across the genome. We show for the first time in a multicellular higher eukaryote that this chromatin landscape is dynamic in response to environmental changes, in this case light or dark growth. Whereas the nucleosome positioning does not change significantly, we show profound changes in the smaller more labile factors under these different conditions. These changes in many cases correspond to the known binding sites of transcription factors that regulate genes in response to light, leading us to propose that full-spectrum chromatin landscape analysis can reveal directly the changes in transcription factor complex binding across the genome.

Introduction

Within the nucleus, higher eukaryotic genomes are packaged in the form of chromatin in which the structural unit is a nucleosome, a ~150 bp of double-stranded DNA wound around an octameric histone protein core. Nucleosomes are arranged in a repetitive manner spaced by a linker region of DNA with a length of ~30 bp, varying by species and genomic location [1,2]. The positioning of nucleosomes relative to DNA sequence is a dynamic process, regulating access to the DNA sequence of various factors and is essential for proper functional biological processes such as transcription, replication, DNA repair and recombination [3]. Moreover, both post-translational modifications of nucleosomal histones and positioning of nucleosomes on the DNA sequence contribute to epigenetic genomic information [4].

Technologies combining digestion of unbound double-stranded DNA using enzymes such as DNase I or micrococcal nuclease (MNase), and next-generation sequencing (NGS) of the remaining protected material allow the generation of genome-wide maps of nucleosome positions. These studies have previously been carried out in a wide range of eukaryotic organisms such as *Saccharomyces cerevisiae* [5], *Caenorhabditis elegans* [6,7], *Drosophila melanogaster* [8], *Zea mays* [9], and also in human cell culture systems [7,10]. Chodavarapu et al. (2010) [11] presented the first nucleosome-positioning map of mononucleosomes in *Arabidopsis thaliana*. They showed that (1) exons are nucleosome-enriched, (2) the intron-exon boundaries are demarcated by strongly positioned flanking nucleosomes, and (3) nucleosomal DNA is methylation-enriched. However their findings focus more on genome region rather than the underlying DNA sequence. Similarly it was shown that genome-wide, nucleosome patterning is uniform in protein-coding genes but not in pseudogenes, transposable element genes and transposable elements [12,13]. Nucleosome patterning is uniform in euchromatin, whereas

pericentromeric heterochromatin shows nucleosome enrichment [11]. Additionally, the periodic distribution of nucleosomes is independent of the level of expression of the gene, but the unoccupied distance between nucleosome positions is higher in transcribed genes [12].

Typically in these studies, digestion of chromatin was performed to completion with optional gel purification steps utilised to enrich for mono-nucleosomal DNA fragments. This gives a limited and static picture of the chromatin landscape focused on the positioning of stable nucleosomes. However, the dynamic structure of chromatin is required for changes in gene expression [14]. Partial digestion of chromatin in yeast and flies reveals variability in nucleosome occupancy profiles explained by the presence of MNase-hypersensitive nucleosomal DNA regions [15–17]. The existence of both hypersensitive and hyper-resistant nucleosomes was shown in the crop plant *Zea mays* [18]. These regions are found in the non-coding DNA around active genes and are colocalised with regulatory elements found in converse non-coding sequences, sequences such as KNOTTED 1 transcription factor binding site [19–21]. Moreover, MNase sensitive regions also correlate with recombination hotspots and hypomethylation [21]. These studies highlight the complex nature of labile DNA-bound proteins.

In addition to the visualisation of nucleosome dynamics, partial digestion approaches allow size-resolution of nuclease protected chromatin particles using modified paired-end mode sequencing protocols to reveal the positions of sub-nucleosome sized particles (subNSPs). These subNSPs protect DNA fragments of a size smaller than 120bp and may represent non-nucleosomal chromatin-associated proteins including DNA replication machinery such as DNA polymerase Origin Recognition complex (ORC) [20] or the transcriptional machinery such as sequence-specific TFs and complexes containing TFs [22]. Further examples include chromatin remodeler complexes [23], and RNA polymerase or proteins involved in chromatin structure [24,25] in DNA recombination [26] or DNA repair [27].

Here we combine partial differential MNase digestion and size-resolved chromatin-seq with transcript profiling from the same samples in *Arabidopsis thaliana* for the first time. Low levels of MNase treatment reveals complexity regarding small particle recruitment and the lability of bound factors which has only previously been suggested in yeast [16] and human sperm [28] through comparisons of mutant types. Here we examine for the first time the dynamics of the overall chromatin landscape in a single cell type to changes in environmental conditions correlated to RNA-seq data, giving insight into the dynamic relationship between chromatin and the changing transcriptome.

Results & discussion

Developing a comprehensive chromatin map of *Arabidopsis thaliana*

Eight replicate *Arabidopsis thaliana* Col-0 cell cultures were sampled after 16-hour passage (see [Methods](#)) (4x Light grown, 4x Dark grown) and subjected to a “higher” or “lower” level of chromatin digestion with MNase followed by Illumina Paired-End sequencing modified to take a wide DNA fragment size range. This allowed us to perform chromatin particle spectrum analysis (CPSA) as previously described in yeast [29] and enabled determination of regions of genomic DNA protected by nucleosomal, multi-nucleosomal, and sub-nucleosomal sized particles (NSP, multiNSP, subNSP) by retaining size information when mapping to genomic DNA.

Spatial mapping of all protected DNA sequences demonstrated a homogeneous coverage of the *A. thaliana* genome with nucleosome sized bound particles, albeit with some regions of lesser chromatin coverage ([Fig 1A](#)). Regions such as the centromeres, the heterochromatic knob on chromosome 4 [30], and the rDNA encoding regions demonstrate abnormal structures likely due to their highly repetitive state inhibiting accurate mapping [31]. Mapping of

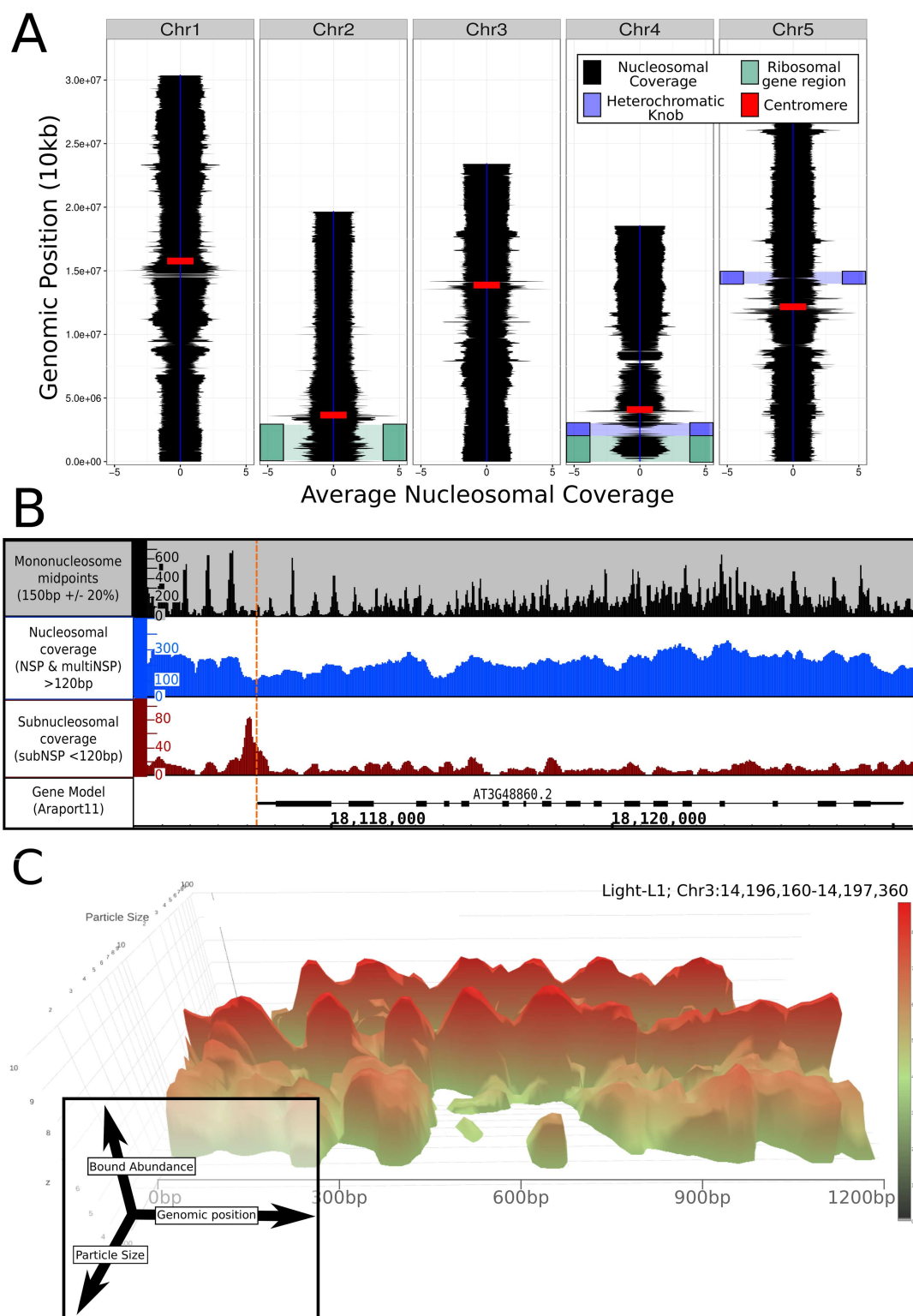


Fig 1. Genomic organisation of chromatin in *Arabidopsis thaliana*. (A) *A. thaliana* genome wide chromatin mapping of nucleosome structure (Mono- and multiNSP binding at 10kb genome regions) with approximate positions of notable genome features [31] demonstrates predominantly homogeneous coverage but with distinct areas of heterogeneous mapping. (B) Gene view demonstrates at organisation of mononucleosomes through mapping of midpoints (black), total mono- and multiNSP genomic protection (blue), and notably, recruitment of a subNSP

(<120bp) at the TSS co-ordinates (red). (C) Genomic protective particles demonstrated continuous size variability (depth-axis) with distinct structure observable at the mono-nucleosomal size, highlighting the complex profile of DNA-bound protein.

<https://doi.org/10.1371/journal.pgen.1006988.g001>

the midpoints of all 150bp (+/-15bp) fragments across the genome (Fig 1B: black track) provides an NSP pattern remarkably consistent with previous results [11,12,32]. However, the approach of mapping according to size provided novel insight into the range of DNA-bound material, as utilising low intensity MNase digestion revealed not only NSPs, but also multiNSP regions resistant to MNase. Coverage of the underlying genome by nucleosomal (~150bp +/-30bp) or multi-nucleosomal ($N * (150\text{bp} + \text{linker})$) chromatin demonstrated a regularly occurring structure throughout the genome, with higher association to genic regions (described further below). Interpolation between sequenced paired reads and cumulative measurement of all mapped fragments equal to or larger than nucleosomal and multi-nucleosomal sized particles generated a single nucleotide resolution map of the total coverage of the underlying DNA at genome scale (Fig 1B: blue track). Furthermore, mapping by particle size in addition to genomic location demonstrated the continuous range of bound particle sizes apparent at this low digestion, as is observed directly in the 1.2kb exemplar region in Fig 1C. This 3D plot shows the non-discrete range of fragment sizes and abundances, revealing the absence of subnucleosomal fragments mapping to its centre, periodic mono-nucleosomes and a range of multi-nucleosomal protected regions.

The differential MNase digestion method was previously implemented in yeast and revealed DNA fragments protected by subNSPs such as TFs [16,29]. To profile the subNSPs across the genome, we computationally subsampled all sequenced material where paired-end reads were less than 120bp i.e. sub-nucleosomal sized regions protected from MNase cutting activity. This revealed that subNSP binding factors have distinct organisation to key positions throughout the genome, specifically positions immediately upstream of, or coinciding with 5' untranslated region (UTR) sites, as observed in the example of Fig 1B (red track). This is consistent with the profiles expected by the binding of TFs and promoter complexes directly upstream of TSSs as observed in yeast [16,29].

Comparison of the otherwise equivalent but differentially digested *A. thaliana* samples between 'high' and 'low' level of MNase treatment demonstrated the effect of digestion extent on the DNA-bound material. At a genomic level, overt variation was not observed in NSP or multiNSP counts between differentially digested samples, consistent with view of chromatin as an essential stabilizer of the genome [5,33] (S1 Fig). However, the marked variation in number of subNSP identified between digestion levels indicates the sensitive nature of these labile factors and suggests that their binding is more readily displaced from the underlying genome than nucleosomes (S1 Fig). Under high digest, subNSPs prominent in low digest samples appear to have a reduced detection indicating greater susceptibility of the underlying DNA to nuclease-mediated degradation due to weaker protection from the bound protein in comparison to nucleosomes. Subsequently, high digest resulted in subNSPs frequently indistinguishable from the likely noise of surrounding low-level small-particle structure, particularly at TSSs (Fig 2A—compare red tracks).

Differential digestion levels in *Zea mays* has previously enabled description of hyper-sensitive sites in the genome where fragile nucleosomes are found [19]. We here observe a similar fragile nucleosome throughout the genome at the '-1' position directly upstream of the TSS (Fig 2B & 2C). Our comparative analysis was consistent with this outcome, which however was not previously reported in *A. thaliana*, likely due to high levels of MNase digestion [12]. For this reason, we believe that the low-digestion level and wider particle size analysis

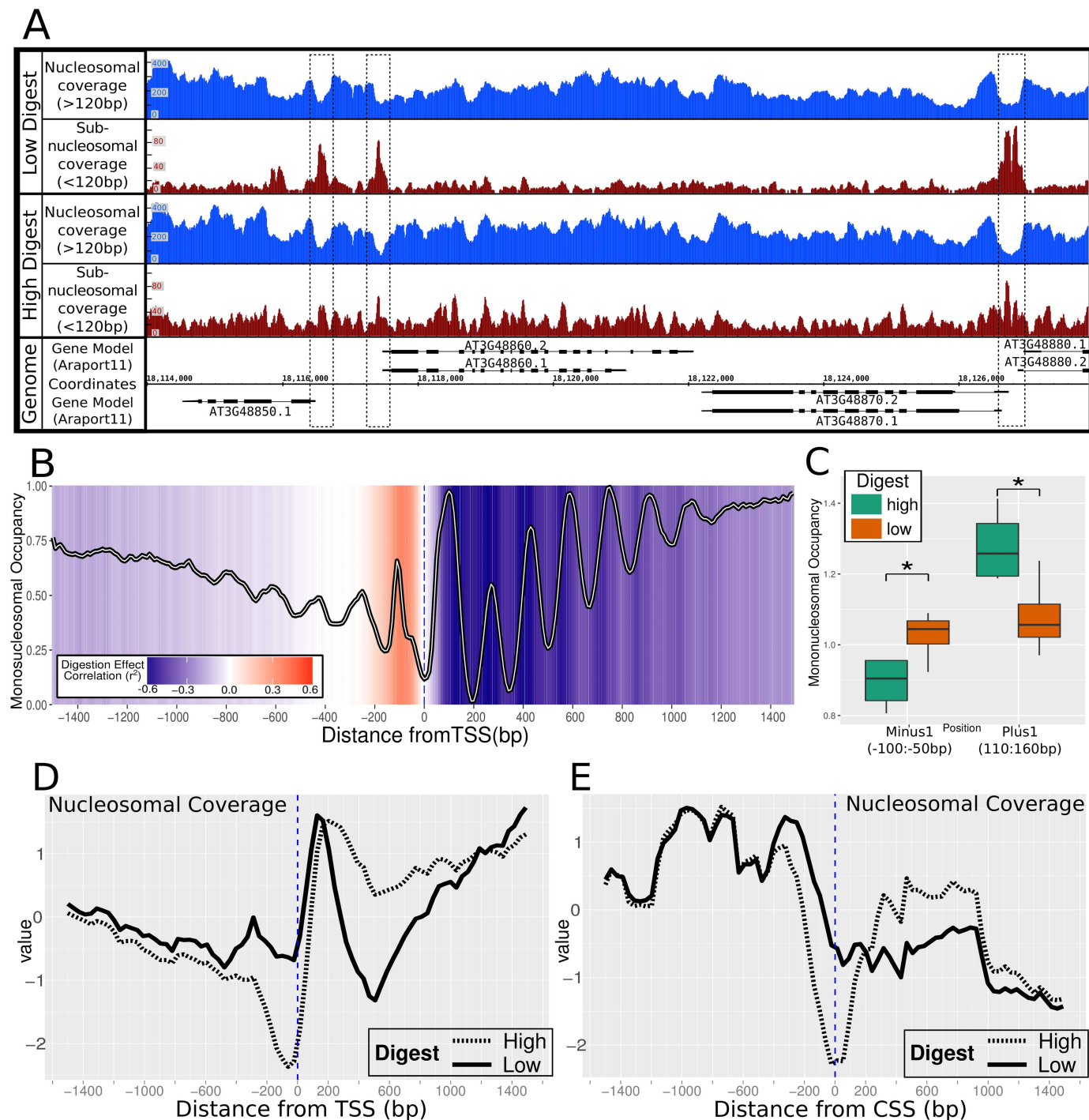


Fig 2. The effect of differential MNase digestion on nucleosomal organisation. (A) NSP and multiNSP coverage occurs throughout the genome but indicates open chromatin at positions of TF recruitment i.e. 5'UTR TSS locations (Blue). Corresponding subNSP recruitment at these locations was markedly less prominent with high-MNase digest preparation (Red). (B) Periodic NSP organisation was conserved at TSS sites (samples = 8, genes = 21,314), with strongly positioned nucleosomes directly downstream from the TSS. Colour scale represents correlation between digestion level estimate and nucleosomal occupancy at that position, where increased digestion correlates with greater monoNSP abundance. (C) Comparison of average occupancy between digestion treatments at 50bp regions in the 'minus-1' and 'plus-1' nucleosome region with significance (* = $P < 0.05$, t.test). Total nucleosomal and multi-nucleosomal coverage surrounding (D) TSS and (E) CSS demonstrates the sensitive labile nucleosomes positioned at the TSS and CSS which are unobserved under high digest.

<https://doi.org/10.1371/journal.pgen.1006988.g002>

implemented here is essential in producing a holistic view of the nucleosomal landscape, and a facet often omitted from other genome-accessibility studies.

Limited digestion reveals varied binding structures in the chromatin landscape of *Arabidopsis* through identification of sub-nucleosomal sized bound particles

The positioning of nucleosomes was observed to be most conserved at the TSS in individual gene examples (Fig 1B). This periodicity is less defined downstream, with the less stringent chromatin organisation being consistent with previous results [12]. Plotting the average structure surrounding the TSS of all *A. thaliana* genes (with consistent expressed isoforms between samples, $N = 21,314$), we observed periodical ~150bp mono-nucleosomal organisation patterning as previously described in *Arabidopsis* (Fig 2B) [11] and typical of other eukaryotic organisms [6,7,17,29]. However, nucleosome positioning in the average profile surrounding the TSS also incorporates the variation between gene lengths and intron-exon structure when surmising multiple gene regions, which contributes to high genic average coverage but lower definition of nucleosomal periodicity. We note the comparative low height of the +2 nucleosome peak which is a consistent feature of our analyses, but currently unexplained.

Consideration of MNase digestion level on mono-nucleosomal organisation shows higher digestion directly correlates with increased mono-nucleosomal structure in the genic region downstream of the TSS. Fig 2B indicates the areas where increased digestion correlates with strengthened or weakened monoNSP positions surmised through the genome, where the colorimetric background shows the strength of correlation between digestion level and NSP abundance per 10bp region. We observe a particularly sensitive nucleosome (negatively correlated to digestion) directly upstream of the TSS which has significantly lower occupancy in the high digested samples (Fig 2C). This is in concordance with the hypersensitive minus-1 nucleosome identified in maize [18]. Additionally, the partially resistant multiNSP and sensitive subNSP positions are less represented under high digest conditions. For nucleosomes this is presumably due to cleavage of di- or tri-nucleosomes to monomers, however the comparative susceptibility of subNSPs results from greater access of MNase to digest under higher levels due to weaker binding or lower levels of DNA protection. The intron-exon boundary demonstrated a strong defining role for nucleosomal positioning as previously described [10], but is not as strong a factor as that by which nucleosomal structure is organised at the TSS (S2 Fig).

Observing the nucleosomal organisation at the TSS regions of individual genes supports variable access of subNSPs due to widening of the open chromatin region between nucleosomes. This pattern is replicated at genomic scale, where open chromatin regions appear at the TSS and Coding Start Sites (CSS) (Fig 2D & 2E), suggesting access for TF recruitment. The differential occupancy signals found at these open regions supports that differential digestion is required to demonstrate the exclusion of the labile and sensitive factors in these regions.

Genome-wide effects of differential growth conditions on chromatin and subNSP structure

The irradiance change in environmental conditions of the diurnal cycle is central to the cyclic expression changes in the plant transcriptome and the normal phenotype [34]. The changing accessibility of key genes has an impact on this expression. Nucleosome spacing at specific genes was correlated with their expression level in targeted analysis [12]. We here addressed the light-responsive gene networks to further understand the role that chromatin has on the activation and inactivation of genes. In addition to NSP binding, we analysed subNSP binding to get a more thorough global view of chromatin dynamics.

Chromatin is understood to be more rigorously structured in exonic regions of genes than intergenic and intronic [11,12,32]. Utilising RNAseq data derived from the same samples used for chromatin-seq, the genome was divided into quartiles based upon gene expression level and the NSP positioning was plotted for four bio-replicates of light grown cell cultures subjected to low and heavy digest (see [Material and methods](#)). Genome-wide, genes under median to high expression showed a strongly focused positioning of NSPs surrounding the TSS ([Fig 3A](#), 50–75%, 75–100%). This structure was less pronounced in low-expressed genes ([Fig 3A](#), 25–50%) and a solitary strongly positioned nucleosome at the point of TSS initiation (up to 200bp downstream) was typical in the lowest quartile of expressed genes ([Fig 3A](#), 0–25%). This loss of mononucleosomal-sized patterning supports the assertion that typical control of gene expression requires well-regulated chromatin [12]. Both the immobile NSP at the TSS of inactive genes, and a strongly defined TSS+1 NSP have been previously reported [12], which was shown to inhibit the binding of transcription complexes and other factors, and to inhibit RNA Pol II activity [35]. Consideration was given to lateral movement of the TSS+1 nucleosome in response to changes in expression, allowing for increased upstream access, however we did not observe convincing evidence for this ([S3 Fig](#)).

We explored the position of subNSPs for the same range of light grown samples. SubNSP binding was enriched at TSS regions throughout the genome. Genes were separated into quartiles by degree of kurtosis of the subNSP peak at the TSS, showing that stronger recruitment correlates with higher levels of expression across the genome ([Fig 3B](#)—upper: example peak structure separation, lower: 4 samples with RNAseq separated by subNSP recruitment). The increased expression in the higher quartiles of subNSP recruitment could either be explained by more frequent binding within the cell population at that position, or stronger binding to DNA by the subNSP, or more defined MNase-hypersensitive linkers at the TSS.

Differential growth conditions effect subNSP changes at single gene resolution

Having established the nucleosomal and sub-nucleosomal structural changes within single samples, we contrasted the chromatin profiles between isolates grown in different irradiance conditions. We did not identify a statistically significant change in nucleosomal patterning genome-wide between samples grown in light or in dark (following two weeks of continuous culture without light). We therefore conclude that the NSP structure is not significantly changed in response to this extrinsic environmental change, even when applied for an extended period. Consequently, we investigated the binding of subNSPs to discern any effect of the environmental changes to the DNA-bound landscape. The correlation between changing gene expression (RNAseq, log₂-fold change) and the recruitment of subNSP (TSS subNSP abundance, log₂-fold change) was assessed for each shared-isoform gene ([Fig 3C](#)). We identified that genes under significantly increased expression in either condition ($p > 0.05$, FDR) correlated with stronger recruitment of subNSP binding at the TSS consistent with DNA-binding complexes such as TF being present at the time of sample harvest.

Hence, whilst no change in gross NSP pattern was seen, changes of expression caused by applied environmental stimuli were correlated with a genome-wide changes at the level of subNSP binding. We therefore sought to understand whether the average landscape observed across the genome was representative of the local profile at well-regulated genes. Within genes of relevance to the light response, we assessed key genes involved in photosynthesis or photosensing that showed strong expression in the RNA-Seq data. For instance, *Photosystem II Subunit S* (CP22, AT1G44575, fold change: -10.8, FDR = 2.59E-136) showed that subNSP recruitment was observed around the TSS and the UTR region directly 5' to the CSS, but was

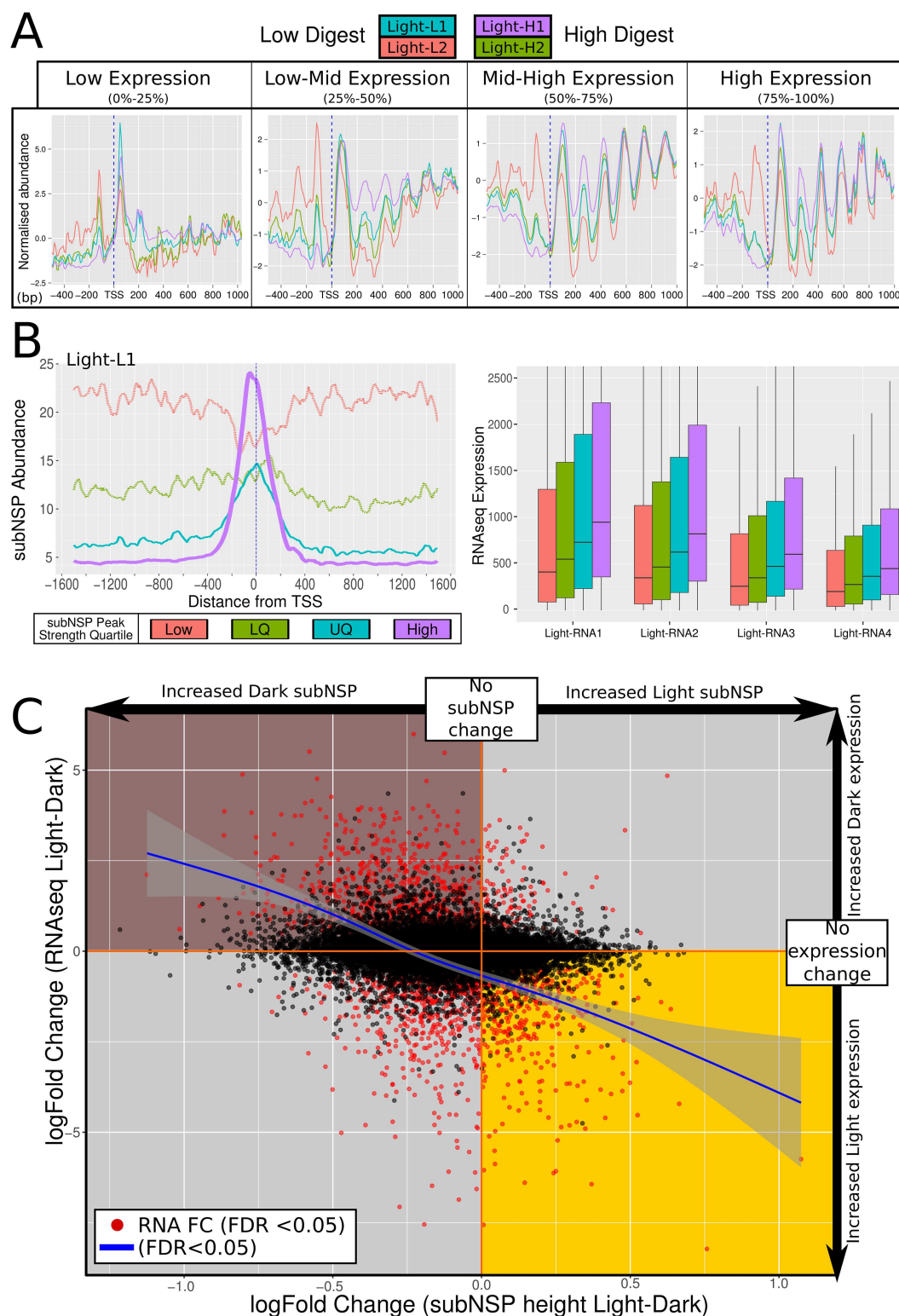


Fig 3. Varying gene expression corresponds to genomic chromatin organisation changes and variation in response to irradiance. (A) NSP structure surrounding TSS separated by quartile of gene expression demonstrated reduced structure correlating with lower expressed genes (light grown samples under high and low digest represented). Inactive and extremely low-expressed genes notably reveal a large +1 nucleosome likely active in inhibiting gene expression. (B) Genes were quartile separated by abundance of subNSP recruitment at TSS as

represented in average profile of quartiles (Light-L1 example: upper). Level of gene expression grouped by quartile demonstrated increasing recruitment and correlated with increased expression within each light-grown sample. (C) Comparison of Light and Dark grown samples for fold change of subNSP abundance was performed against corresponding gene expression. Genes under significant expression change (FDR<0.05, red) demonstrated greater recruitment of subNSP in accordance with higher expression (Trendline: loess smoothed, 0.95 CI: grey boundary).

<https://doi.org/10.1371/journal.pgen.1006988.g003>

not identified in samples grown in dark (Fig 4A, dashed boxes). *CP22* encodes a photoprotective pigment-binding protein associated with the photosystem II in the grana thylakoids [36]. To prevent the formation of reactive species and the damage to membranes, a quenching of excited chlorophyll occurs by transferring the energy to carotenoid pigments [37]. Another example is the disappearance of a 5'UTR associated subNSP in dark compared to light for the *Rubisco Small Subunit 3B* (AT5G38410, fold change: -9.46, FDR = 3.01E-076, Fig 4B), encoding a subunit of the key Calvin cycle enzyme ribulose 1,5-bisphosphate carboxylase-oxygenase (RuBisCO) [38]. Again, the presence of subNSP at the 5'UTR appears to coincide reliably with the increased expression.

We then examined whether these changes in subNSPs could be related to known positions of TF binding to specific genomic locations. Previously identified TF Binding Sites (TFBS) were sourced from Athamap [39,40] (S4 Fig) or O'Malley et. al (2016) [41] to assess subNSP binding to their corresponding genomic targets. *PHYTOCHROME INTERACTING FACTOR 3* (PIF3) is a helix-loop-helix TF interacting with photoreceptors phyA and phyB as part of the plant light response [42]. Analysis of the PIF3 target binding motif across the genome demonstrated significantly more occupancy in light conditions (Fig 5A, Number of sites = 1,930). This was despite a non-significant change (log2FC: -0.25, FDR = 0.7) in expression of the PIF3 gene, consistent with post-translational regulation of the TF through phosphorylation [43]. The related PIF4 TF was however associated with a significant downregulation in dark-grown cells (AT2G43010, log2FC: -1.4, FDR = 0.079) and revealed a similar subNSP binding structure at identified sites (Fig 5B, N = 20,252)[44].

In contrast to the light responsive PIF TFs, the Myb-type *CIRCADIAN CLOCK ASSOCIATED* (CCA1, AT2G46830) TF similarly demonstrated increased binding at known TFBS [45] in light-grown samples, but presents an alternative saddle-shaped binding pattern. This heightened accumulated around the binding site indicates the motif region having susceptibility to MNase digestion and the subNSP being offset in dark conditions (Fig 5C, N = 59,249). TF encoding genes with higher expression in dark grown samples (ERF1, WRKY59, AT4G18450) did not demonstrate an increased abundance of subNSP binding and patterning was retained, although some variability can be observed due to variation in digestion level (S5 Fig).

While the majority of subNSP binding positions are associated with genic promotor regions, approximately 7.91% of positions identified were greater than 10kb from an annotated coding sequence and 1.18% greater than 100kb distant. These sites can clearly be observed as protected DNA fragments of unknown function, which supports the role of the subNSP/TF detection approach as a new methodology to reveal not only the potential binding of transcription factors but also indicating other regions of interest for further analysis. However, while these positions may function as binding locations for long distance promotor elements, there are also potential roles unrelated to transcription factor binding which must be considered including the spatial organisation of DNA, and underlying physical structure of DNA which make regions resistant to MNase digestion i.e. secondary structures. Furthermore, investigation into associations with underlying genomic factors and histone modification statuses will be of significant future interest due to the connection of modified histones to transcriptional

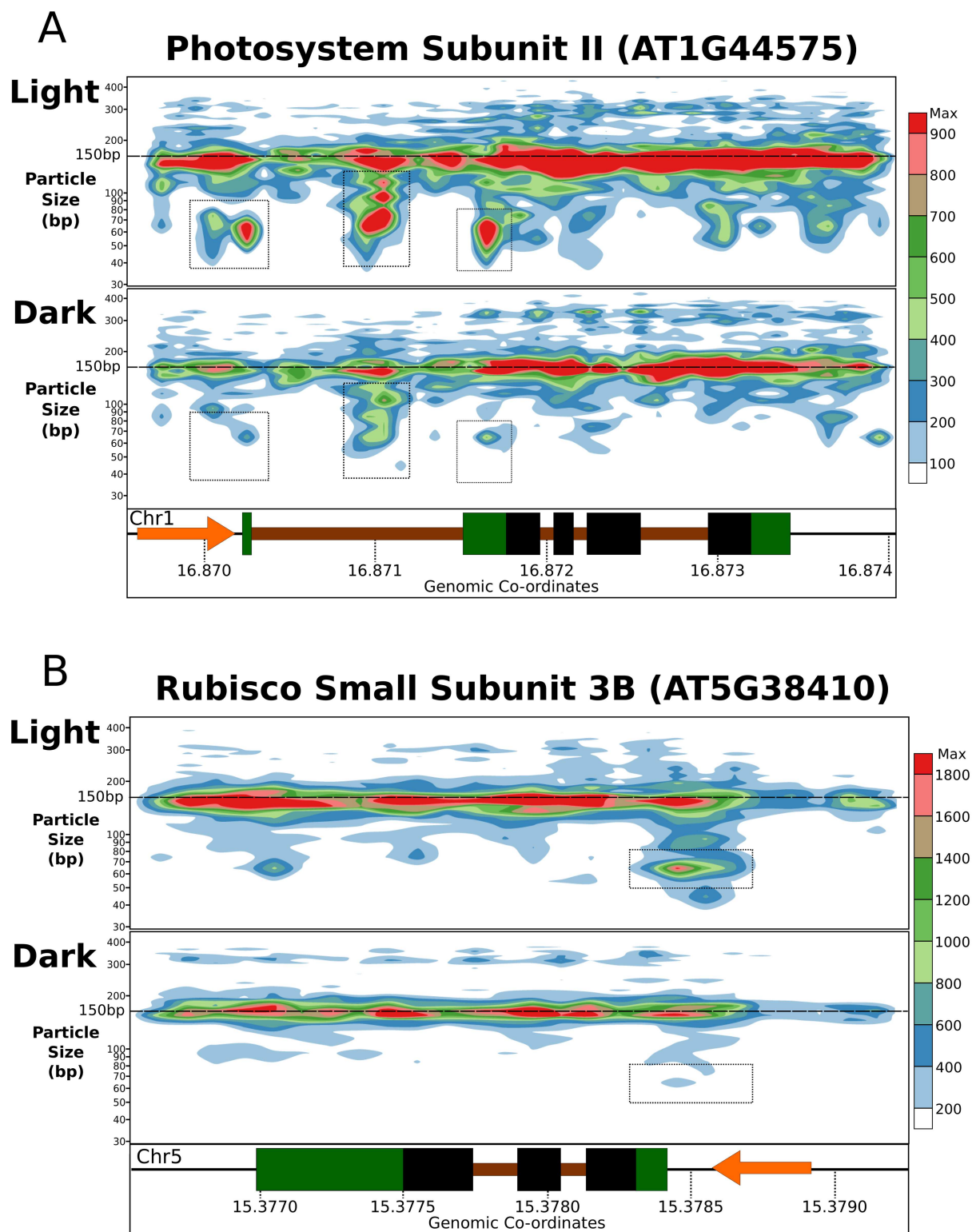


Fig 4. SubNSP recruitment occurs preferentially at active genes. Total coverage mapping of all sized DNA-binding particles for low-digested samples (Light-L1/L2 and Dark-L1/L2) demonstrates recruitment of subNSPs to specific TSS and UTR positions in the key light-response genes (A) Photosystem Subunit II (NPQ4, AT1G44575) and (B) Rubisco Small Subunit 3B (RBCS3B, AT5G38410). Observed subNSP recruitment is absent from dark-grown samples which are almost totally inactive (PSII—FC: -10.8, FDR = 2.59E-136, RBCS3B —FC: -9.46, FDR = 3.01E-076). Colour scale represents abundance of mapped fragments normalised for sequencing depth.

<https://doi.org/10.1371/journal.pgen.1006988.g004>

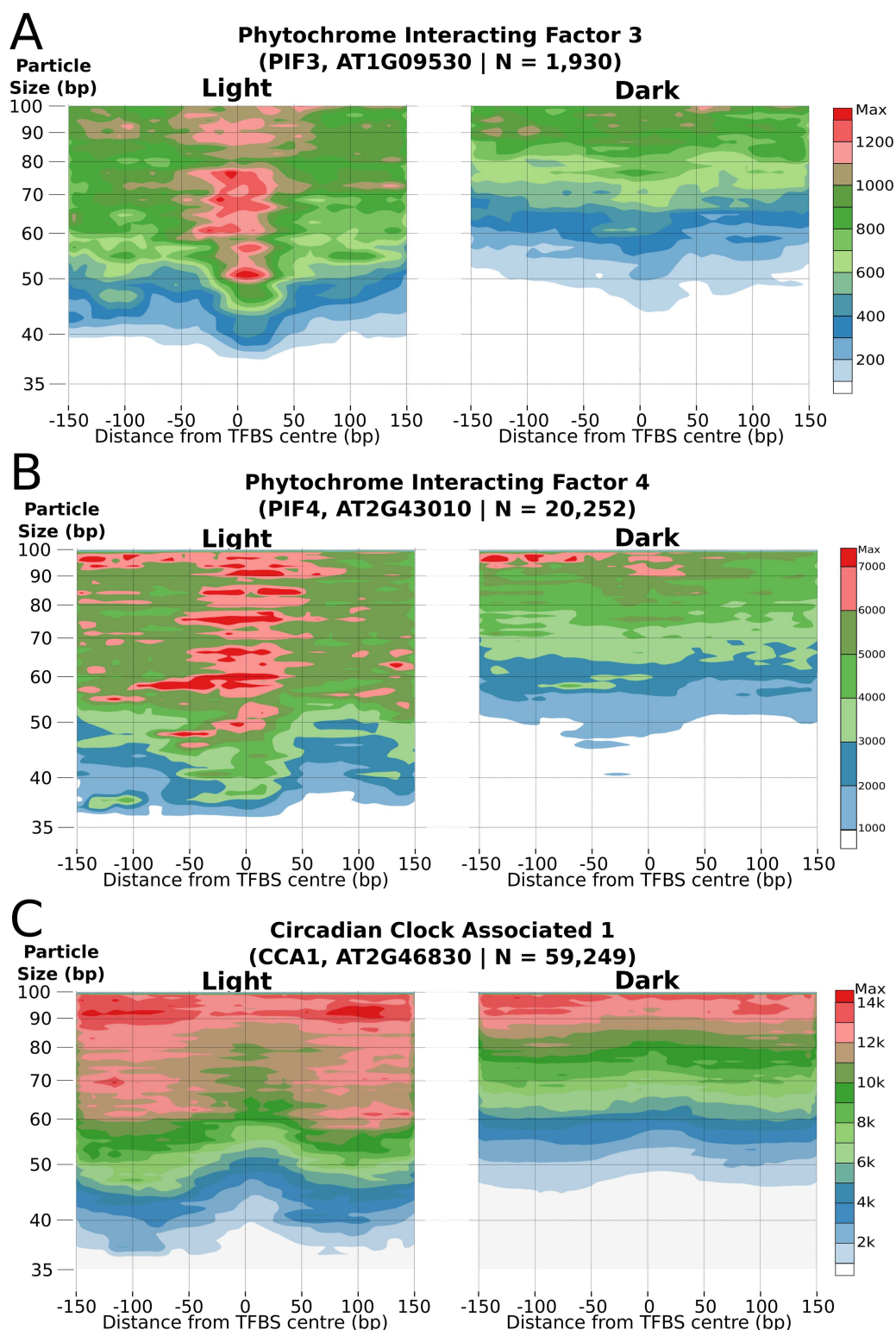


Fig 5. SubNSP organisation at transcription factor binding sites. TFBS positions were assayed for subNSP binding for light-responsive Transcription Factors (A) PIF3 (n = 1,930), (B) PIF4 (n = 20,252) and (C) the CCA1 (N = 59,249). Direct visualisation highlights the differential binding of the subNSP across the genome between growth conditions and identifying the TF response to the extrinsic irradiance changes.

<https://doi.org/10.1371/journal.pgen.1006988.g005>

control [46]. One aspect to be assessed under matched conditions is the connecting the exclusion of H2A.Z and methylation of DNA in actively transcribed DNA [47], and positions found to be enriched for subNSP binding ie. the 5'UTR and Transcription Start Sites. As ChIP-seq approaches to mapping TFs in *A. thaliana* advance, we should come to learn the identities of the factors we have observed protecting DNA and their role in the genomic landscape.

Materials & methods

Condition for cultivation of *Arabidopsis thaliana* cell line

Dispersed cell suspension cultures, (prepared from *Arabidopsis thaliana* leaves Columbia eco-type, Col-0 [48,49] and were a kind gif of Dr. Linda Hanley-Bowdoin obtained from North Carolina State University, USA) were used. The cultures are a homogeneous population of physiologically and morphologically identical cells [50,51]. The cultures were maintained in 250 mL Erlenmeyer flasks, filled with 50 mL of cell culture medium (Gamborg's B5 basal medium with minor organic (Sigma G5893) in 1.1 mg/L 2,4-D, 3 mM MES and 3% sucrose) as previously described [49,52]. Cells were grown on a rotary shaker 160 rpm at 23°C (LS-X (Lab Shaker), Kuhner Shaker X). Constant light was used whereas dark grown cells were incubated under the same shaking and temperature condition but the flask were covered of aluminium foil. Cell line was subcultured every 7 days with a 2:50 (inoculum: fresh medium) dilution ratio. Cultures were carried out in duplicates.

Cell sampling

Light grown cells were sampled 5 days or 16 hours after subculture. Dark grown cells were adapted to dark conditions for 2 weeks with a 7-day subculture period. Dark grown cells were sampled 16-hour after the 2nd subculture. Cells were sampled by harvesting 30mL of cell culture and washed with autoclaved ultrapure water. Cells were frozen in liquid nitrogen and ground into a fine powder. The cell powder was either stored at -80°C or used directly for MNase digestion and RNA isolation.

Chromatin digestion

Pelleted cell culture was ground in a liquid nitrogen-cooled pestle and mortar to generate a cell powder. 1mL of the powder was suspended in a 0.5mL modified spheroplast digestion buffer & Nonidet P40 (SDBN: 1M sorbitol, 10mM NaCl, 50 mM Tris-HCl pH7.5, 5mM MgCl₂, 1 mM CaCl₂, 1 mM 2-mercaptoethanol, 0.5 mM spermidine, 0.075% Nonidet P40). 300 µl of cells were transferred into 1.5mL microcentrifuge tube containing 30 or 120 units of MNase (Affymetrix). Cells were digested with MNase at 37°C for 3 min. The MNase digestion reactions were stopped by addition of and thorough mixing 30 µl of STOP solution containing 5% SDS and 250 µl EDTA. DNA was extracted by addition of an equal volume of 2X CTAB supplemented with PVP (200 mM Tris-HCl, 40 mM EDTA, pH 8.0; 2.8 M NaCl, 4% w/v CTAB, 2% PVP-40 CAS number 9003-39-8) was added to the digestion reaction and incubated at 45°C for 15 mins. DNA was separated from protein by two phenol: chloroform (1:1; 600µl total) steps. DNA was precipitated with sodium acetate and propan-2-ol, washed in 80% ethanol and dried. Samples were incubated with 10X RNase A at 37°C for 1h with 100U unmodified T4 polynucleotide kinase (NEB) for further 30 min at 37°C.

MNase digested DNA was separated on 1.5% agarose gel, stained with ethidium bromide for 10 min at 80V, DNA fragments ranging from 20 bp to 1 kb were excised and gel pieces containing DNA were placed in a Costar Spin-X 0.45 µm cellulose acetate centrifuge tube filter (Sigma CLS8136). Two series of freeze-thaws (-80C for 10 mins/RT for 10 mins) were

performed to macerate the gel. Filter tubes were centrifuged twice (14kg; 5 min; RT) with a 180 degree rotation of the tubes between the two spins. DNA in the aqueous phase was then extracted with 400µl phenol: chloroform (1:1), and precipitated at -80°C for 30 mins with sodium acetate to propan-2-ol and finally washed with 80% ethanol before being resuspended in ultrapure milliQ water.

RNA isolation

Total RNA was extracted using the RNeasy Plant mini Kit (Qiagen, <http://www.qiagen.com>).

Nucleic acid quantification and quality check

DNA samples were quantified on Qubit™ 2.0 Fluorometer with dsDNA BR assay and RNA with Qubit RNA BR assay according to manufacturer's instructions. RNA quality was checked on 1.5% agarose gel.

Library preparation and sequencing

Exeter Sequencing Service and Computational core facilities at the University of Exeter performed the Genomic and RNA library preparation, quantification and 50bp paired end sequencing on HiSeq 2500. Additional information can be found on the Exeter Sequencing Service website (<http://sequencing.exeter.ac.uk/>). A total of ~861M paired end reads were obtained from the MNase-seq genomic library and ~173M for the RNAseq library (S6 Fig).

Data processing and bioinformatic analysis

Genomic mapping was performed with bowtie allowing for 3 nucleotide errors [53], resulting in 85.09% alignment to the TAIR10 reference Columbia-0 genome [54]. Aligned BAM files were converted to wig trace files for monoNSP midpoints (150bp +/- 10%) using bespoke perl scripts. DNA coverage was produced using bamToBed and coverageBed from the bedtools package [55]. Datasets were either complete (total) or where specified split into <120bp, and >120bp subsamples, and converted to wig traces (S7 Fig). Gene feature alignments were produced using danpos [56] and the ARAPORT11 annotation [57]. Where alternate splice variants were present between samples as determined from RNAseq analysis (see below) only genes primarily sharing the same model were utilised to ensure accurate boundary comparisons i.e. Transcription Start Sites.

Continuous size variation plots (2D and 3D) were produced using a bespoke package (BAM2SizePlot.py) producing visualisations utilising paired-end insert distance to determine particle size and can be obtained at <https://github.com/ChromatinCardiff/ALD>.

RNAseq mapping was performed with Tophat2 [58], resulting in 78.40% alignment to the TAIR10 reference Colombia-0 genome [54] and the ARAPORT11 annotation [57]. Determination of isoform presence was achieved with cufflinks [59] and only primary isoforms shared between all samples were utilised as gene feature boundaries to ensure error was not induced between splice variants (n = 21,314). Quantitative assessment was performed with HTseq [60] and analysed with edgeR [61,62].

Supporting information

S1 Fig. Table: Nucleosomal and subNSP abundance throughout the *Arabidopsis thaliana* genome.
(XLSX)

S2 Fig. Mono-nucleosomal coverage at intron-exon boundary.
(TIFF)

S3 Fig. Boxplot demonstrating differential distance of +1 nucleosome from TSS between Light and dark grown samples in correlation to change in expression.
(TIFF)

S4 Fig. Athamap motif matrices predicting TFBS for PIF3 [42], PIF4 [44] and CCA1 [45].
(TIFF)

S5 Fig. TF encoding genes with higher expression in dark grown samples did not demonstrate differential patterning at TF binding sites.
(TIFF)

S6 Fig. Table: MNase-seq and RNAseq read counts and mapping success.
(XLSX)

S7 Fig. Total Coverage (TC) quantification method and gene example.
(TIFF)

Acknowledgments

We wish to thank Joanne Kilby for technical support, Harley Worthy for initial data analysis and the Cardiff School of Biosciences Genome Research Hub.

Author Contributions

Conceptualization: Nicholas A. Kent, James A. H. Murray.

Formal analysis: Daniel Antony Pass, Nicholas A. Kent.

Funding acquisition: Nicholas A. Kent, James A. H. Murray.

Investigation: Daniel Antony Pass, Emily Sornay, Angela Marchbank, Margaret R. Crawford, Konrad Paszkiewicz, Nicholas A. Kent, James A. H. Murray.

Methodology: Daniel Antony Pass, Emily Sornay, Angela Marchbank, Margaret R. Crawford, Konrad Paszkiewicz, Nicholas A. Kent, James A. H. Murray.

Project administration: Nicholas A. Kent, James A. H. Murray.

Resources: Nicholas A. Kent, James A. H. Murray.

Software: Daniel Antony Pass.

Supervision: Angela Marchbank, Nicholas A. Kent, James A. H. Murray.

Validation: Emily Sornay, Angela Marchbank, Margaret R. Crawford, Konrad Paszkiewicz, Nicholas A. Kent.

Visualization: Daniel Antony Pass, Nicholas A. Kent.

Writing – original draft: Daniel Antony Pass, Emily Sornay.

Writing – review & editing: Daniel Antony Pass, Emily Sornay, Nicholas A. Kent, James A. H. Murray.

References

1. Szerlong HJ, Hansen JC. Function To Dynamic Chromatin Structure. 2011; 89: 24–34.

2. Bradbury EM, K. E. Van Holde. Chromatin Series in molecular biology. Journal of Molecular Recognition. New York: Springer-Verlag; 1989.
3. Kornberg RD. Chromatin Structure: A Repeating Unit of Histones and DNA. Science (80-). 1974; 184: 868–871. <https://doi.org/10.1126/science.184.4139.868>
4. Khorasanizadeh S. The Nucleosome. Cell. 2004; 116: 259–272. [https://doi.org/10.1016/S0092-8674\(04\)00044-3](https://doi.org/10.1016/S0092-8674(04)00044-3) PMID: [14744436](#)
5. Kaplan N, Moore IK, Fondufe-Mittendorf Y, Gossett AJ, Tillo D, Field Y, et al. The DNA-encoded nucleosome organization of a eukaryotic genome. Nature. Nature Publishing Group; 2009; 458: 362–366. <https://doi.org/10.1038/nature07667> PMID: [19092803](#)
6. Andersson R, Enroth S, Rada-Iglesias A, Wadelius C, Komorowski J. Nucleosomes are well positioned in exons and carry characteristic histone modifications. Genome Res. 2009; 19: 1732–1741. <https://doi.org/10.1101/gr.092353.109> PMID: [19687145](#)
7. Tilgner H, Nikolaou C, Althammer S, Sammeth M, Beato M, Valcárcel J, et al. Nucleosome positioning as a determinant of exon recognition. Nat Struct Mol Biol. Nature Publishing Group; 2009; 16: 996–1001. <https://doi.org/10.1038/nsmb.1658> PMID: [19684599](#)
8. Mavrich TN, Jiang C, Ioshikhes IP, Li X, Venters BJ, Zanton SJ, et al. Nucleosome organization in the *Drosophila* genome. Nature. 2008; 453: 358–362. <https://doi.org/10.1038/nature06929> PMID: [18408708](#)
9. Fincher JA, Vera DL, Hughes DD, McGinnis KM, Dennis JH, Bass HW. Genome-Wide Prediction of Nucleosome Occupancy in Maize Reveals Plant Chromatin Structural Features at Genes and Other Elements at Multiple Scales. PLANT Physiol. 2013; 162: 1127–1141. <https://doi.org/10.1104/pp.113.216432> PMID: [23572549](#)
10. Schwartz S, Meshorer E, Ast G. Chromatin organization marks exon-intron structure. Nat Struct Mol Biol. Nature Publishing Group; 2009; 16: 990–995. <https://doi.org/10.1038/nsmb.1659> PMID: [19684600](#)
11. Chodavarapu RK, Feng S, Bernatavichute Y V, Chen P-YY, Stroud H, Yu Y, et al. Relationship between nucleosome positioning and DNA methylation. Nature. 2010; 466: 388–392. <https://doi.org/10.1038/nature09147> PMID: [20512117](#)
12. Li G, Liu S, Wang J, He J, Huang H, Zhang Y, et al. ISWI proteins participate in the genome-wide nucleosome distribution in *Arabidopsis*. Plant J. 2014; 78: 706–714. <https://doi.org/10.1111/tpj.12499> PMID: [24606212](#)
13. Singh M, Bag S, Bhardwaj A, Ranjan A, Mantri S, Nigam D, et al. Global nucleosome positioning regulates salicylic acid mediated transcription in *Arabidopsis thaliana*. BMC Plant Biol. 2015; 15: 13. <https://doi.org/10.1186/s12870-014-0404-2> PMID: [25604550](#)
14. Cairns BR. The logic of chromatin architecture and remodelling at promoters. Nature. 2009; 461: 193–198. <https://doi.org/10.1038/nature08450> PMID: [19741699](#)
15. Xi Y, Yao J, Chen R, Li W, He X. Nucleosome fragility reveals novel functional states of chromatin and poises genes for activation. Genome Res. 2011; 21: 718–724. <https://doi.org/10.1101/gr.117101.110> PMID: [21363969](#)
16. Henikoff JG, Belsky JA, Krassovsky K, MacAlpine DM, Henikoff S. Epigenome characterization at single base-pair resolution. Proc Natl Acad Sci U S A. 2011; 108: 18318–18323. <https://doi.org/10.1073/pnas.1110731108> PMID: [22025700](#)
17. Chereji R V., Kan TW, Grudniewska MK, Romashchenko A V., Berezikov E, Zhimulev IF, et al. Genome-wide profiling of nucleosome sensitivity and chromatin accessibility in *Drosophila melanogaster*. Nucleic Acids Res. 2015; 44: 1036–1051. <https://doi.org/10.1093/nar/gkv978> PMID: [26429969](#)
18. Vera DL, Madzima TF, Labonne JD, Alam MP, Hoffman GG, Girmurugan SB, et al. Differential nucleosome sensitivity profiling of chromatin reveals biochemical footprints coupled to gene expression and functional DNA elements in maize. Plant Cell. 2014; 26: 3883–93. <https://doi.org/10.1105/tpc.114.130609> PMID: [25361955](#)
19. Vera DL, Madzima TF, Labonne JD, Alam MP, Hoffman GG, Girmurugan SB, et al. Differential Nucleosome Sensitivity Profiling of Chromatin Reveals Biochemical Footprints Coupled to Gene Expression and Functional DNA Elements in Maize. 2014; 26: 3883–3893. <https://doi.org/10.1105/tpc.114.130609> PMID: [25361955](#)
20. Bass HW, Wear EE, Lee TJ, Hoffman GG, Gumber HK, Allen GC, et al. A maize root tip system to study DNA replication programmes in somatic and endocycling nuclei during plant development. J Exp Bot. 2014; 65: 2747–2756. <https://doi.org/10.1093/jxb/ert470> PMID: [24449386](#)
21. Rodgers-Melnick E, Vera DL, Bass HW, Buckler ES. Open chromatin reveals the functional maize genome. Proc Natl Acad Sci. 2016; 113: E3177–E3184. <https://doi.org/10.1073/pnas.1525244113> PMID: [27185945](#)

22. Kaufmann K, Pajoro A, Angenent GC. Regulation of transcription in plants: mechanisms controlling developmental switches. *Nat Rev Genet*. Nature Publishing Group; 2010; 11: 830–842. <https://doi.org/10.1038/nrg2885> PMID: [21063441](#)
23. Brzeski J, Jerzmanowski A. Plant chromatin—Epigenetics linked to ATP-dependent remodeling and architectural proteins. *FEBS Lett*. 2004; 567: 15–19. <https://doi.org/10.1016/j.febslet.2004.03.068> PMID: [15165887](#)
24. Pedersen DS, Merkle T, Marktl B, Lildballe DL, Antosch M, Bergmann T, et al. Nucleocytoplasmic Distribution of the Arabidopsis Chromatin-Associated HMGB2/3 and HMGB4 Proteins. *Plant Physiol*. 2010; 154: 1831–1841. <https://doi.org/10.1104/pp.110.163055> PMID: [20940346](#)
25. Wang H, Dittmer T a, Richards EJ. Arabidopsis CROWDED NUCLEI (CRWN) proteins are required for nuclear size control and heterochromatin organization. *BMC Plant Biol*. 2013; 13: 200. <https://doi.org/10.1186/1471-2229-13-200> PMID: [24308514](#)
26. Knoll A, Fauser F, Puchta H. DNA recombination in somatic plant cells: Mechanisms and evolutionary consequences. *Chromosom Res*. 2014; 22: 191–201. <https://doi.org/10.1007/s10577-014-9415-y> PMID: [24788060](#)
27. Manova V, Gruszka D. DNA damage and repair in plants—from models to crops. *Front Plant Sci*. 2015; 6: 1–26.
28. Carone BR, Hung JH, Hainer SJ, Chou M Te, Carone DM, Weng Z, et al. High-resolution mapping of chromatin packaging in mouse embryonic stem cells and sperm. *Dev Cell*. Elsevier Inc.; 2014; 30: 11–22. <https://doi.org/10.1016/j.devcel.2014.05.024> PMID: [24998598](#)
29. Kent NA, Adams S, Moorhouse A, Paszkiewicz K. Chromatin particle spectrum analysis: A method for comparative chromatin structure analysis using paired-end mode next-generation DNA sequencing. *Nucleic Acids Res*. 2011; 39. <https://doi.org/10.1093/nar/gkq1183> PMID: [21131275](#)
30. Fransz PF, Armstrong S, De Jong JH, Parnell LD, Drunen C Van, Dean C, et al. Integrated Cytogenetic Map of Chromosome Arm 4S of *A. thaliana*: Structural Organization of Heterochromatic Knob and Centromere Region. 2000; 100: 367–376.
31. The Arabidopsis Genome Initiative. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature*. 2000; 408: 796–815. <https://doi.org/10.1038/35048692> PMID: [11130711](#)
32. Morton T, Petricka J, Corcoran DL, Li S, Winter CM, Carda a., et al. Paired-End Analysis of Transcription Start Sites in *Arabidopsis* Reveals Plant-Specific Promoter Signatures. *Plant Cell*. 2014; 26: tpc.114.125617-. <https://doi.org/10.1105/tpc.114.125617> PMID: [25035402](#)
33. Van Bortle K, Corces VG. The role of chromatin insulators in nuclear architecture and genome function. *Curr Opin Genet Dev*. 2013; 23: 212–218. <https://doi.org/10.1016/j.gde.2012.11.003> PMID: [23298659](#)
34. Nolte C, Staiger D. RNA around the clock—regulation at the RNA level in biological timing. *Front Plant Sci*. 2015; 6: 1–15.
35. Weber C, Ramachandran S, Henikoff S. Nucleosomes are context-specific, H2A.Z-Modulated barriers to RNA polymerase. *Mol Cell*. Elsevier Inc.; 2014; 53: 819–830. <https://doi.org/10.1016/j.molcel.2014.02.014> PMID: [24606920](#)
36. Caffarri S, Tibiletti T, Jennings RC, Santabarbara S. A comparison between plant photosystem I and photosystem II architecture and functioning. *Curr Protein Pept Sci*. 2014; 15: 296–331. <https://doi.org/10.2174/1389203715666140327102218> PMID: [24678674](#)
37. Correa-Galvis V, Poschmann G, Melzer M, Stühler K, Jahns P. PsbS interactions involved in the activation of energy dissipation in *Arabidopsis*. *Nat Plants*. 2016; 2: 15225. <https://doi.org/10.1038/nplants.2015.225> PMID: [27249196](#)
38. Carmo-Silva E, Scales JC, Madgwick PJ, Parry MAJ. Optimizing Rubisco and its regulation for greater resource use efficiency. *Plant, Cell Environ*. 2015; 38: 1817–1832. <https://doi.org/10.1111/pce.12425> PMID: [25123951](#)
39. Hehl R, Norval L, Romanov A, Bu L, Braunschweig D, Email C. Boosting AthaMap Database Content with Data from Protein Binding Microarrays Special Online Collection—Database Paper. 2016; 57: 2–6.
40. Steffens NO, Galuschka C, Schindler M, Bülow L, Hehl R. AthaMap: an online resource for in silico transcription factor binding sites in the *Arabidopsis thaliana* genome. *Nucleic Acids Res*. 2004; 32: D368–72. <https://doi.org/10.1093/nar/gkh017> PMID: [14681436](#)
41. O'Malley RC, Huang S shan C, Song L, Lewsey MG, Bartlett A, Nery JR, et al. Erratum: Cistrome and Epicistrome Features Shape the Regulatory DNA Landscape (Cell (2016) 165(5) (1280???1292)). *Cell*. Elsevier Inc.; 2016; 166: 1598.
42. Martínez-García JF, Huq E, Quail PH. Direct Targeting of Light Signals to a Promoter Element-Bound Transcription Factor. *Science* (80-). 2000; 288: 859–863. <https://doi.org/10.1126/science.288.5467.859>
43. Kim J, Shen Y, Han Y, Park J, Kirchenbauer D, Soh M, et al. Phytochrome Phosphorylation Modulates Light Signaling by Influencing the Protein—Protein Interaction. 2004; 16: 2629–2640.

44. Pruneda-Paz JL, Breton G, Nagel DH, Kang SE, Bonaldi K, Doherty CJ, et al. A Genome-Scale Resource for the Functional Characterization of Arabidopsis Transcription Factors. *Cell Rep*. 2014; 8: 622–632. <https://doi.org/10.1016/j.celrep.2014.06.033> PMID: [25043187](#)
45. Franco-Zorrilla JM, López-Vidriero I, Carrasco JL, Godoy M, Vera P, Solano R. DNA-binding specificities of plant transcription factors and their potential to define target genes. *Proc Natl Acad Sci U S A*. 2014; 111: 2367–72. <https://doi.org/10.1073/pnas.1316278111> PMID: [24477691](#)
46. Roudier F, Ahmed I, Bérard C, Sarazin A, Mary-Huard T, Cortijo S, et al. Integrative epigenomic mapping defines four main chromatin states in Arabidopsis. *EMBO J*. 2011; 30: 1928–1938. <https://doi.org/10.1038/emboj.2011.103> PMID: [21487388](#)
47. Zilberman D, Coleman-Derr D, Ballinger T, Henikoff S. Histone H2A.Z and DNA methylation are mutually antagonistic chromatin marks. *Nature*. 2008; 456: 125–9. <https://doi.org/10.1038/nature07324> PMID: [18815594](#)
48. Lee T, Pascuzzi PE, Settlege SB, Shultz RW, Tanurdzic M, Pablo D, et al. Arabidopsis thaliana Chromosome 4 Replicates in Two Phases That Correlate with Chromatin State. 2010; 6. <https://doi.org/10.1371/journal.pgen.1000982> PMID: [20548960](#)
49. Tanurdzic M, Vaughn MW, Jiang H, Lee T-J, Slotkin RK, Sosinski B, et al. Epigenomic consequences of immortalized plant cell suspension culture. *PLoS Biol*. 2008; 6: 2880–2895. <https://doi.org/10.1371/journal.pbio.0060302> PMID: [19071958](#)
50. Sequeira-Mendes J, Aragón I, Peiró R, Mendez-Giraldez R, Zhang X, Jacobsen SE, et al. The Functional Topography of the Arabidopsis Genome Is Organized in a Reduced Number of Linear Motifs of Chromatin States. *Plant Cell*. 2014; 26: 2351–2366. <https://doi.org/10.1105/tpc.114.124578> PMID: [24934173](#)
51. Seifertová D, Klíma P, Paľvrezová M, Petrášek J, Zažímalová E, Opatrný Z. Plant Cell Lines in Cell Morphogenesis Research. In: Žárský V, Cvrčková F, editors. *Plant Cell Morphogenesis: Methods and Protocols*. Totowa, NJ: Humana Press; 2014. pp. 215–229. https://doi.org/10.1007/978-1-62703-643-6_18 PMID: [24132432](#)
52. Lee TJ, Pascuzzi PE, Settlege SB, Shultz RW, Tanurdzic M, Rabinowicz PD, et al. Arabidopsis thaliana chromosome 4 replicates in two phases that correlate with chromatin state. *PLoS Genet*. 2010; 6: 1–18. <https://doi.org/10.1371/journal.pgen.1000982> PMID: [20548960](#)
53. Langmead B, Trapnell C, Pop M, Salzberg S. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol*. 2009; 10: R25. <https://doi.org/10.1186/gb-2009-10-3-r25> PMID: [19261174](#)
54. Swarbreck D, Wilks C, Lamesch P, Berardini TZ, Garcia-Hernandez M, Foerster H, et al. The Arabidopsis Information Resource (TAIR): Gene structure and function annotation. *Nucleic Acids Res*. 2008; 36: 1009–1014.
55. Quinlan AR, Hall IM. BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics*. 2010; 26: 841–842. <https://doi.org/10.1093/bioinformatics/btq033> PMID: [20110278](#)
56. Chen K, Xi Y, Pan X, Li Z, Kaestner K, Tyler J, et al. DANPOS: Dynamic analysis of nucleosome position and occupancy by sequencing. *Genome Res*. 2013; 23: 341–351. <https://doi.org/10.1101/gr.142067.112> PMID: [23193179](#)
57. Cheng C-Y, Krishnakumar V, Chan A, Schobel S, Town CD. Araport11: a complete reannotation of the Arabidopsis thaliana reference genome. *bioRxiv*. 2016; 47308. <https://doi.org/10.1101/047308>
58. Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol*. 2013; 14: R36. <https://doi.org/10.1186/gb-2013-14-4-r36> PMID: [23618408](#)
59. Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, et al. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat Protoc*. 2012; 7: 562–78. <https://doi.org/10.1038/nprot.2012.016> PMID: [22383036](#)
60. Anders S, Pyl PT, Huber W. HTSeq-A Python framework to work with high-throughput sequencing data. *Bioinformatics*. 2015; 31: 166–169. <https://doi.org/10.1093/bioinformatics/btu638> PMID: [25260700](#)
61. Robinson MD, McCarthy DJ, Smyth GK. edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*. 2009; 26: 139–140. <https://doi.org/10.1093/bioinformatics/btp616> PMID: [19910308](#)
62. McCarthy DJ, Chen Y, Smyth GK. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res*. 2012; 40: 4288–4297. <https://doi.org/10.1093/nar/gks042> PMID: [22287627](#)